

1. `Adult`: True or False
2. `belongs_to_collection`: '{"id': 10194, 'name': 'Toy Story Collection', 'poster_path': '/7G9915LfUQ2lVfwMEEhDsn3kT4B.jpg', 'backdrop_path': '/9FBWqcd9IRuEDUrTdcaafOMKUq.jpg}'"
3. `budget`: Movie budget
4. `genres`: "[{'id': 16, 'name': 'Animation'}, {'id': 35, 'name': 'Comedy'}, {'id': 10751, 'name': 'Family'}]"
5. `homepage`: 'http://toystory.disney.com/toy-story'
6. `id`: Movie ID
7. `imdb_id`: IMDB IDtt0114709
8. `original_language`: 92 languages
9. `original_title`: original title of the movie
10. `overview`: comments or feedback
11. `popularity`: popularity value 0 to 66
12. `poster_path`: /rhIRbceoE9IR4veEXuwCC2wARtG.jpg
13. `production_companies`: [{'name': 'Pixar Animation Studios', 'id': 3}]
14. `production_countries`: [{'iso_3166_1': 'US', 'name': 'United States of America'}]
15. `release_date`: 1995-10-30
- 16.
17. `Revenue`: revenue generated by the movie
18. `Runtime`: movie length
19. `spoken_languages`: [{'iso_639_1': 'en', 'name': 'English'}]
20. `status`: ['Released', nan, 'Rumored', 'Post Production', 'In Production', 'Planned', 'Canceled']
21. `tagline`: Roll the dice and unleash the excitement!
22. `Title`: movie title
23. `Video`: True / False
24. `vote_average`: 0 -10
25. `vote_count`: 0-14075

`data size` 45466, 24

Simple recommended systems



$$\left(\frac{v}{v+m} \times R \right) + \left(\frac{m}{v+m} \times C \right)$$

- v is the number of votes garnered by the movie
- m is the minimum number of votes required for the movie to be in the chart (the prerequisite)
- R is the mean rating of the movie
- C is the mean rating of all the movies in the dataset

Knowledge based recommended systems

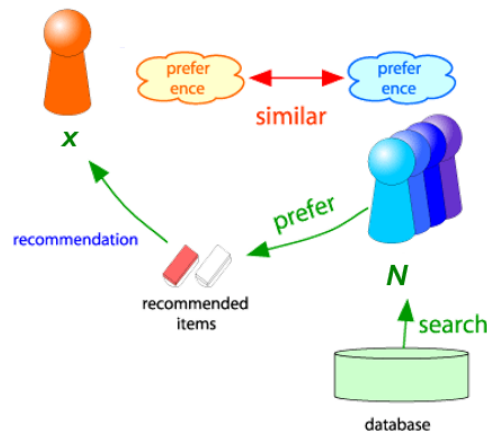


1. Ask the user for the genres of movies he/she is looking for
2. Ask the user for the duration
3. Ask the user for the timeline of the movies recommended
4. Using the information collected, recommend movies to the user that have a high weighted rating (according to the IMDB formula) and that satisfy the preceding conditions

14

Collaborative Filtering

- Consider user x
- Find set N of other users whose ratings are “**similar**” to x ’s ratings
- Estimate x ’s ratings based on ratings of users in N



Evaluation

movies

users

Test Data Set

1	3	4			
	3	5			5
		4	5		5
		3			
		3			
2			?		?
				?	
	2	1			?
	3			?	
1					